Poisson Disk Sampling on the Grassmannian: Applications in Subspace Optimization

Bhavya Kailkhura

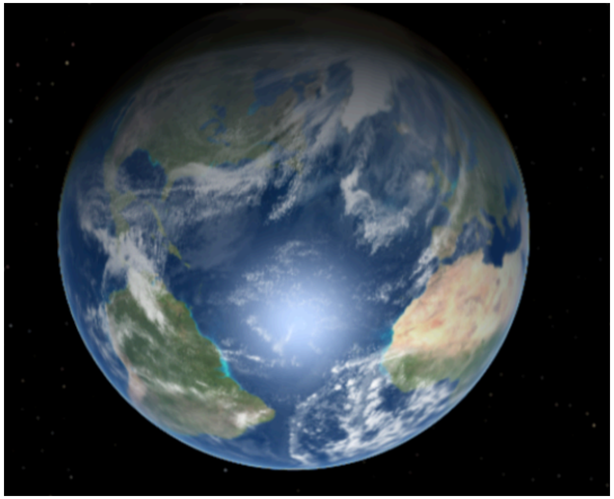
Collaborators: Rushil Anirudh, Jayaraman Thiagarajan, Timo Bremer

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Lawrence Livermore National Laboratory

Non-Euclidean Manifolds in Computer Vision

- Why should we care? Because, we cannot ignore the real geometry of the mathematical spaces.



What is the distance between Paris and San Francisco?

- An N-dimensional **manifold** M is a topological space where every point is endowed with “local” Euclidean structure

Optimization on Non-Linear Manifolds

- *Matrix Lie Group:* Set of *n x n* non-singular matrices with a smooth manifold structure

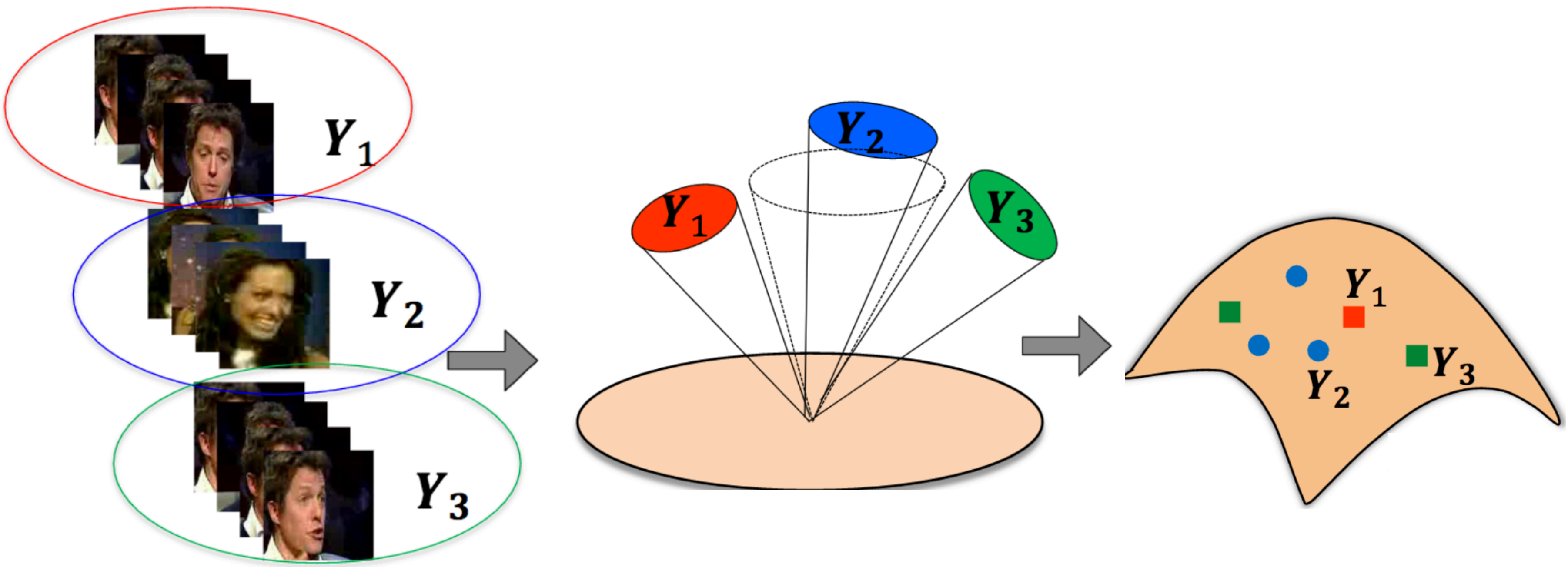
$$\mathcal{SO}(n) = \left\{ Y \in \mathbb{R}^{n \times n} : Y^T Y = I,\ det(Y) = 1 \right\}$$

A truck rendered at different orientations on *SO(2)* – Rotation Matrices

- *Stiefel Manifold:* Set of orthonormal matrices of size *n x p*

- *Grassmann Manifold:* Set of p-dimensional linear subspaces in *n*-dimensions

Subspace Analysis on Grassmann Manifolds



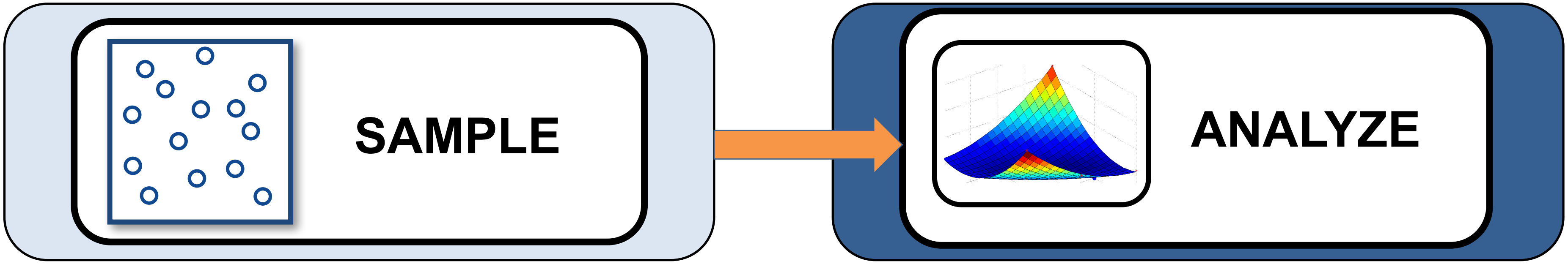
Videos can be treated as image sets

Each set modeled as a linear subspace

Analysis on the Grassmann manifold

Gradient descent based optimization requires understanding of the characteristics of loss functions on non-Euclidean manifolds

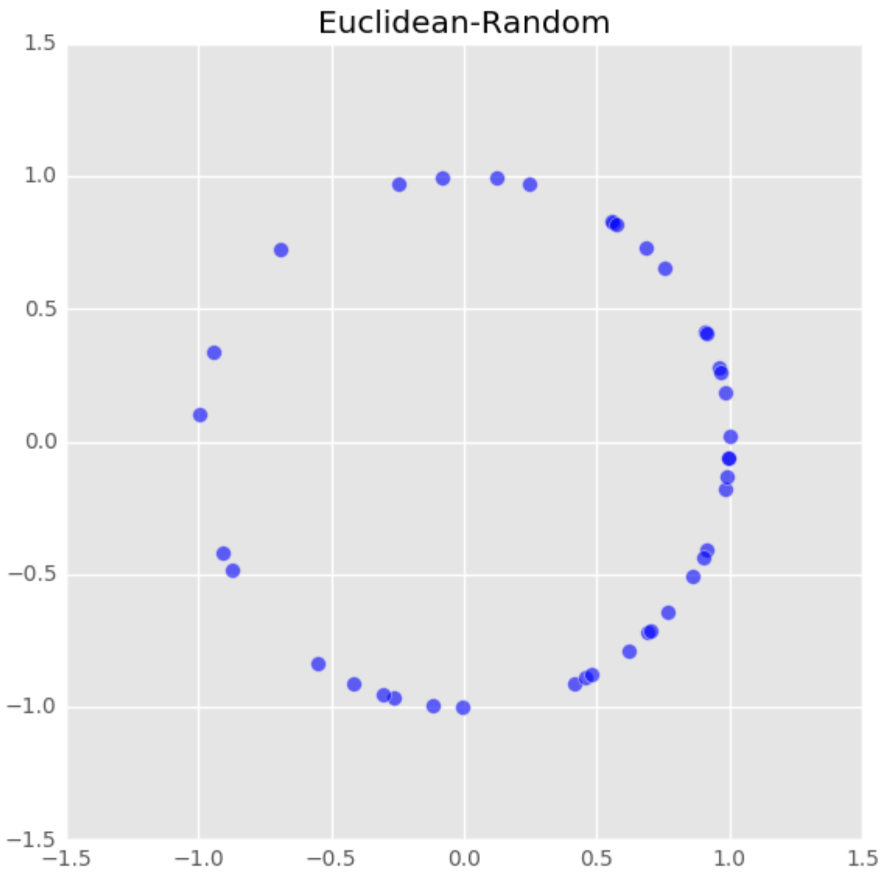
A Statistical Approach

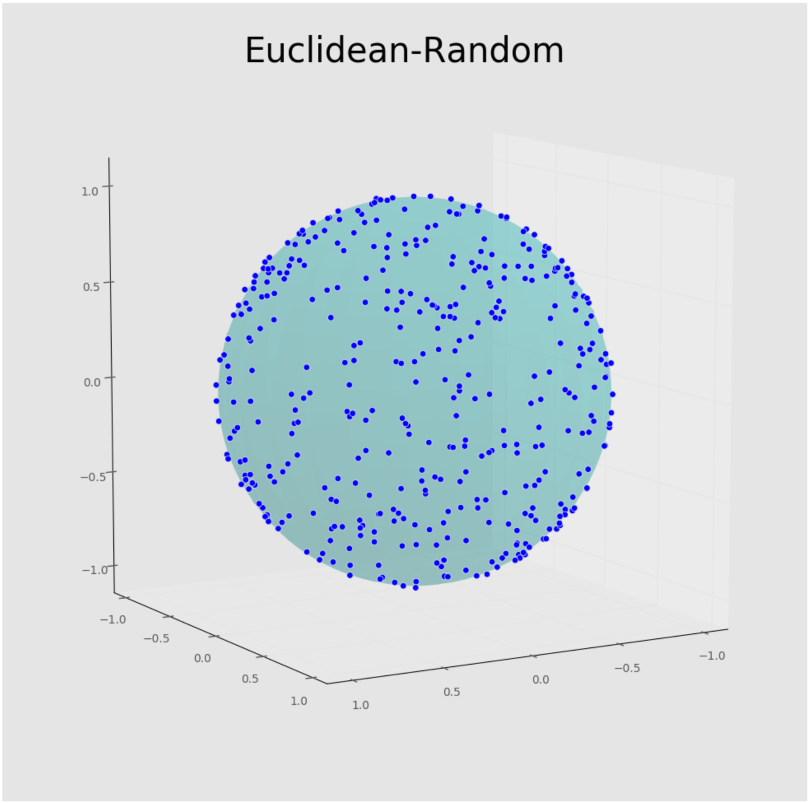


Improved Sample Designs Can Enable a Better Understanding of the Optimization Surface

Existing Techniques Provides Poor Coverage

Projecting random samples created in the Euclidean space results in poor coverage of the manifold





Characteristics of a Good Sampling Pattern

- In its most generic form the inference problem can be described as recovering or analyzing a multi-variate, smooth function

$$f : \mathcal{M} \mapsto \mathbb{R}^d$$

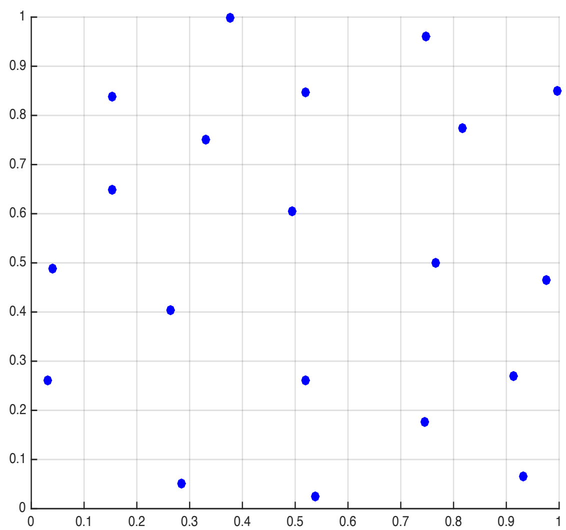
defined on some manifold $\mathcal{M} \subset \mathbb{R}^n$.

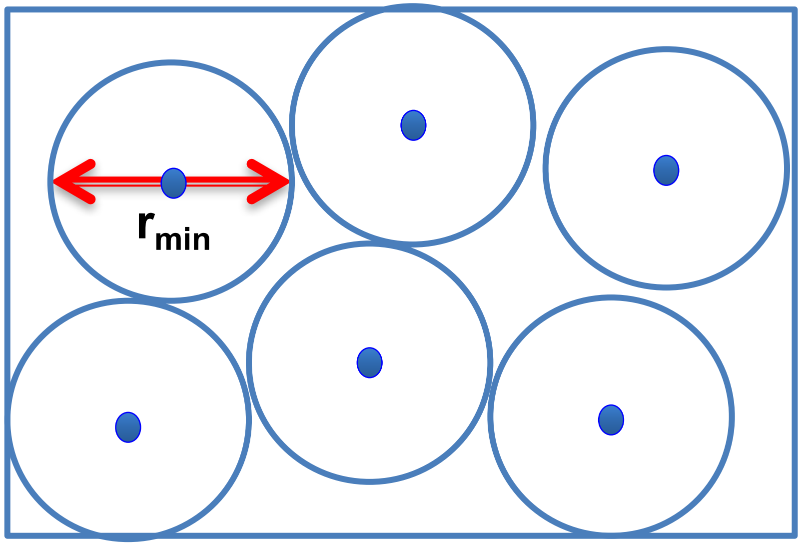
- A good sampling pattern is expected to have two main properties
  1. the sampling should be random
  2. the samples should uniformly cover all the manifold

Sampling Euclidean Spaces: Poisson Disk Sampling

Generating Poisson Disk Samples

- Each sample is placed with uniform probability
- No two samples are closer than a pre-specified radius
- A PDS is maximal when no more points can be inserted





Grassmann Manifolds

- Linear subspaces of dimension *k* in *n*-dimensional Euclidean spaces are points on the Grassmannian *Gr*(*n*,*k*)

Intrinsic Distance on the Grassmannian

Given two subspaces $\mathbf{A}, \mathbf{B} \in Gr(n,k)$, the geodesic distance between them is measured as $d(\mathbf{A}, \mathbf{B}) = (\sum_{i=1}^{k} \theta_i^2)^{1/2}$, where $\{\theta_i\}_{i=1}^{k}$ are the principal angles.

- Measures the smallest rotation that takes from one subspace to another
- Chordal distance: $(\sum_{i=1}^{k} \sin^2 \theta_i)^{1/2} = \frac{1}{\sqrt{2}} \|\mathbf{A}\mathbf{A}^T - \mathbf{B}\mathbf{B}^T\|_F.$

Dart throwing for Poisson Disk Sampling

Algorithm 1 Dart Throwing on the Grassmann manifold

Require: Dimensions (n, k), number of samples N and r_{min}, $\mathbf{S} = \emptyset$

1: **while** $|\mathbf{S}| \leq N$ **do**

2: Throw a Dart i:

- Generate random matrix $\mathbf{Z}_i \in \mathbb{R}^{n \times k}$
- Obtain the corresponding point $Q \in \mathcal{G}_{n,k}$ as the QR decomposition of $\mathbf{Z}_i$.
- Assign $\mathbf{S}_i \leftarrow Q$

3: **if** $d_{\mathcal{G}}(\mathbf{S}_i, \mathbf{S}_j) \leq r_{min}$, $\forall \mathbf{S}_j \in \mathbf{S}$ **then**

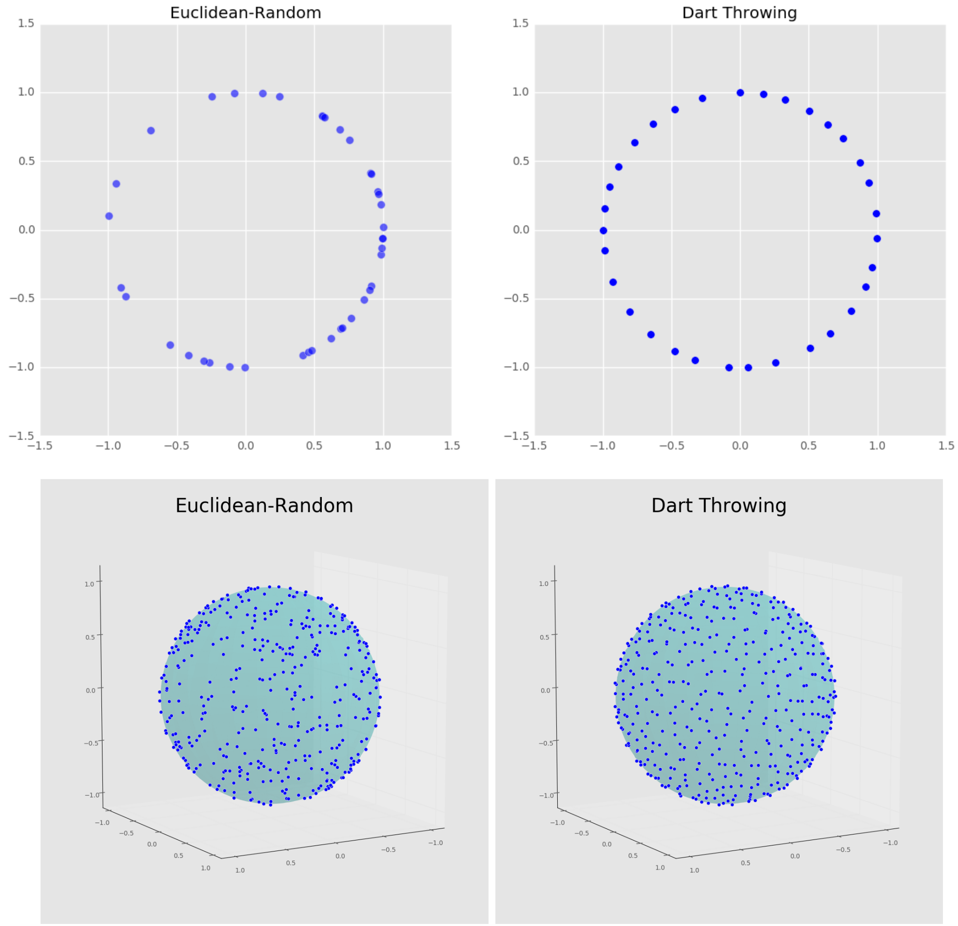
4: Add sample $\mathbf{S}_i$ to the point set $\mathbf{S}$

5: **end if**

6: **end while**

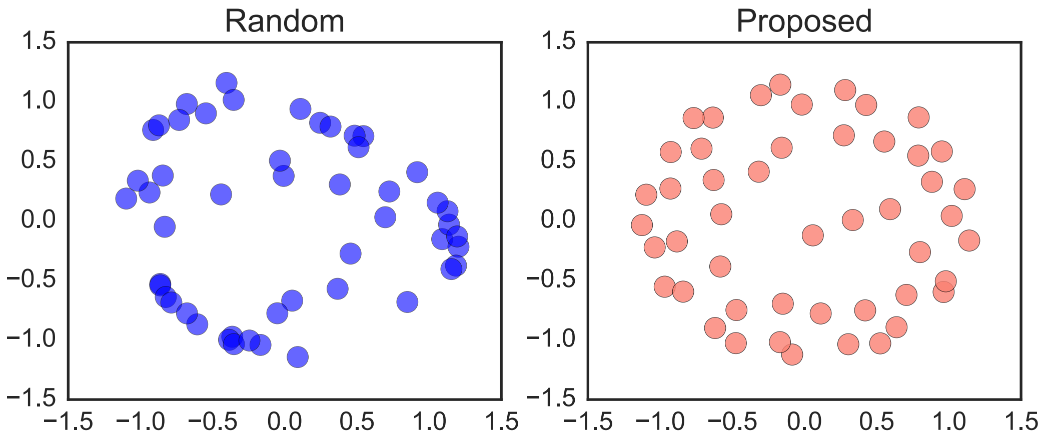
Proposed Scheme Provides Better Coverage

- Our proposed approach produces better uniform samples, thereby leading to improved analysis and optimization

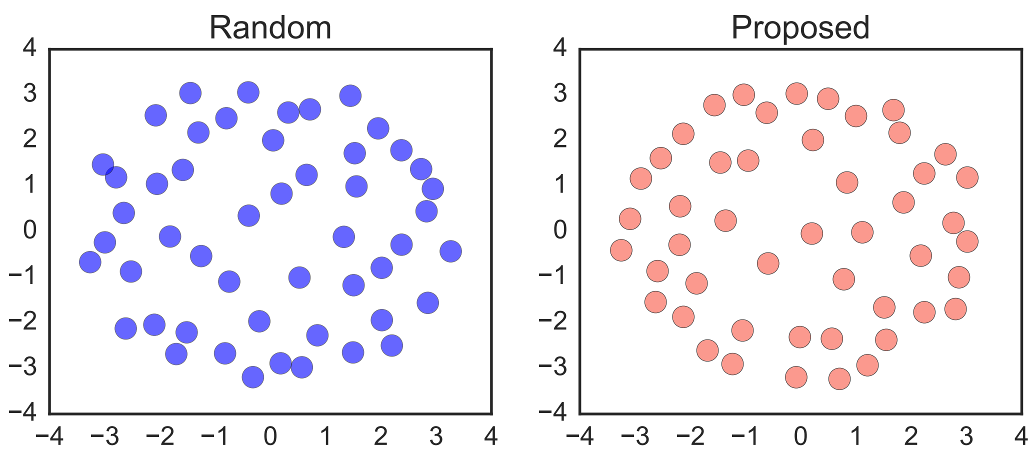


Visualizing the sample distributions (MDS)

50 samples from G_3^2



50 samples from G_8^4



Subspace Optimization using Grassmann Samples

- Subspace Learning as Graph Embedding

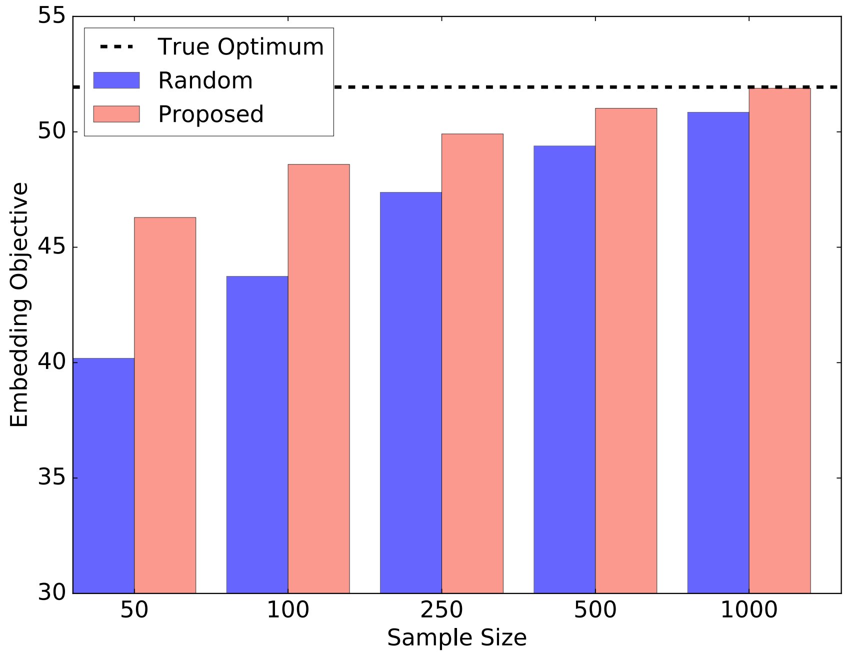
$$\mathbf{Y}^* = arg \min_{tr(\mathbf{V}^T\mathbf{X}\mathbf{B}\mathbf{X}^T\mathbf{V})=d} tr(\mathbf{V}^T\mathbf{X}\mathbf{L}\mathbf{X}^T\mathbf{V})$$

- Solution of PCA/LDA is approximated as a linear combination of the Dart throwing samples $\mathbf{S}_\mathbf{i}$
- Consensus based approach

$$\mathbf{V}^* = arg \min_{\mathbf{V}^T\mathbf{V}=\mathbf{I}} tr\left(\mathbf{V}^T\sum_{i=1}^{N}\left(\alpha_i\mathbf{I} - \alpha_i\mathbf{S}_i\mathbf{S}_i^T\right)\mathbf{V}\right)$$

PDS Samples are Very Effective in Preserving the Desired Relationships

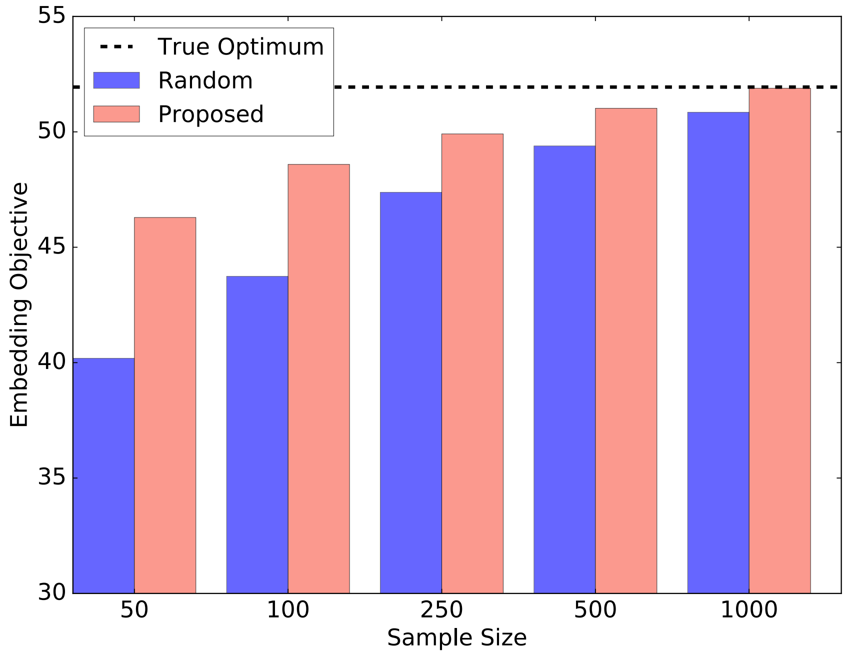
- Principal Component Analysis cost optimization using consensus on the Grassmannian (higher is better)



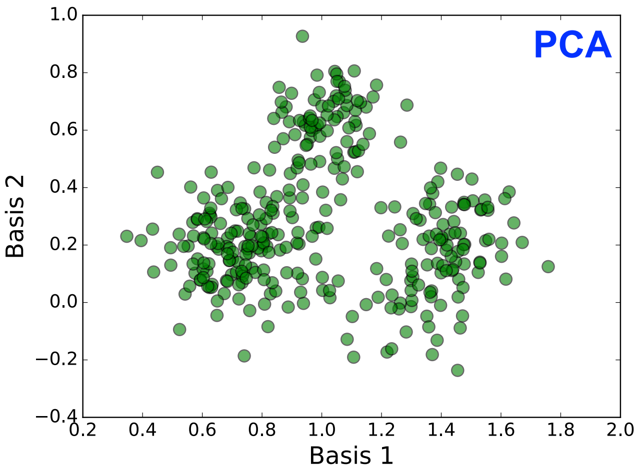
Ecoli Dataset (7 dimensional)

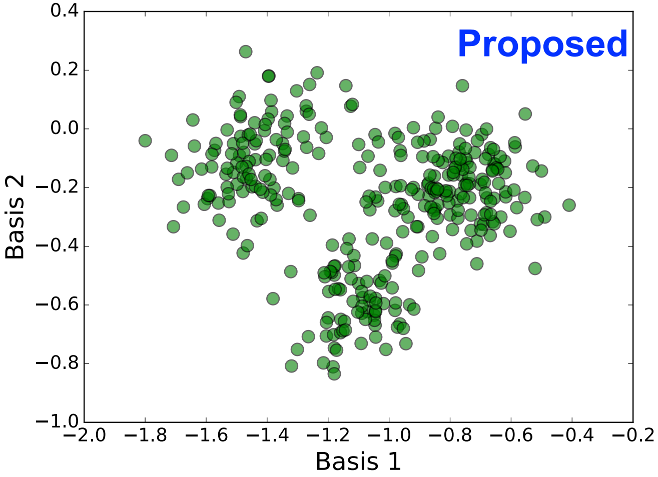
PDS Samples are Very Effective in Preserving the Desired Relationships

- Principal Component Analysis cost optimization using consensus on the Grassmannian (higher is better)



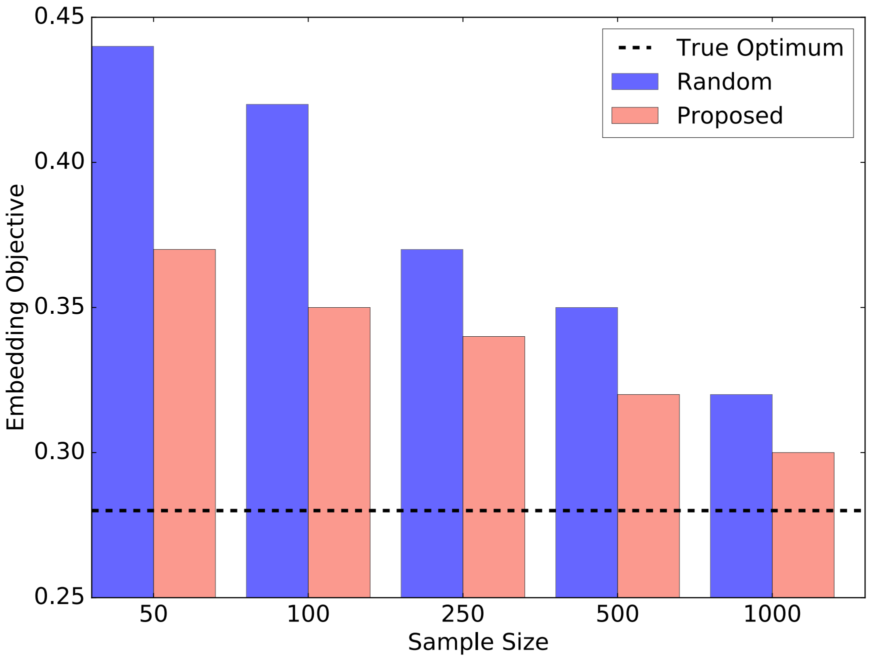
Ecoli Dataset (7 dimensional)





PDS Samples are Very Effective in Preserving the Desired Relationships

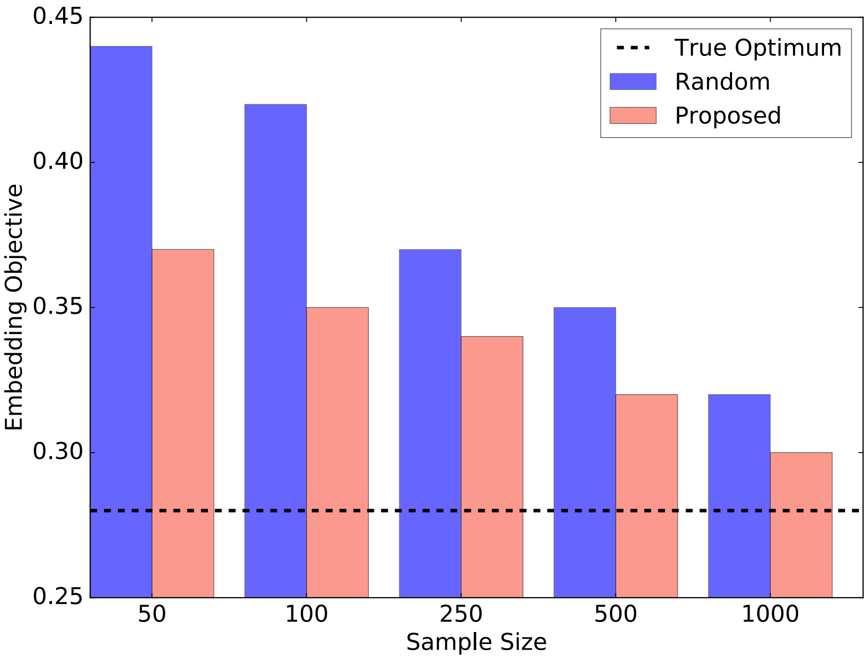
- Comparison of the consensus embedding obtained using the LDA cost on the Grassmannian, to the true solution obtained using generalized eigenvalue decomposition



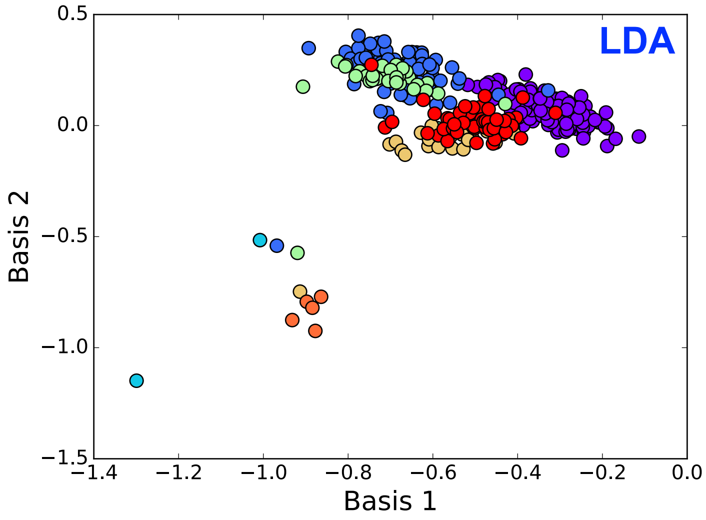
Ecoli Dataset (7 dimensional)

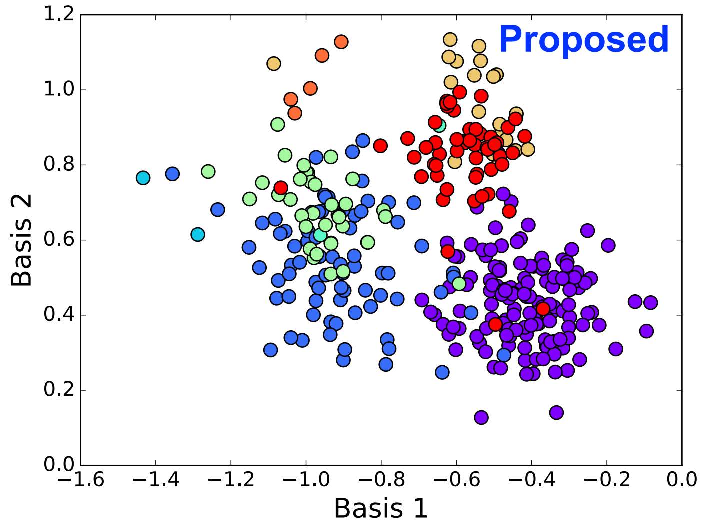
PDS Samples are Very Effective in Preserving the Desired Relationships

- Comparison of the consensus embeddin cost on the Grassmannian, to the true so generalized eigenvalue decomposition



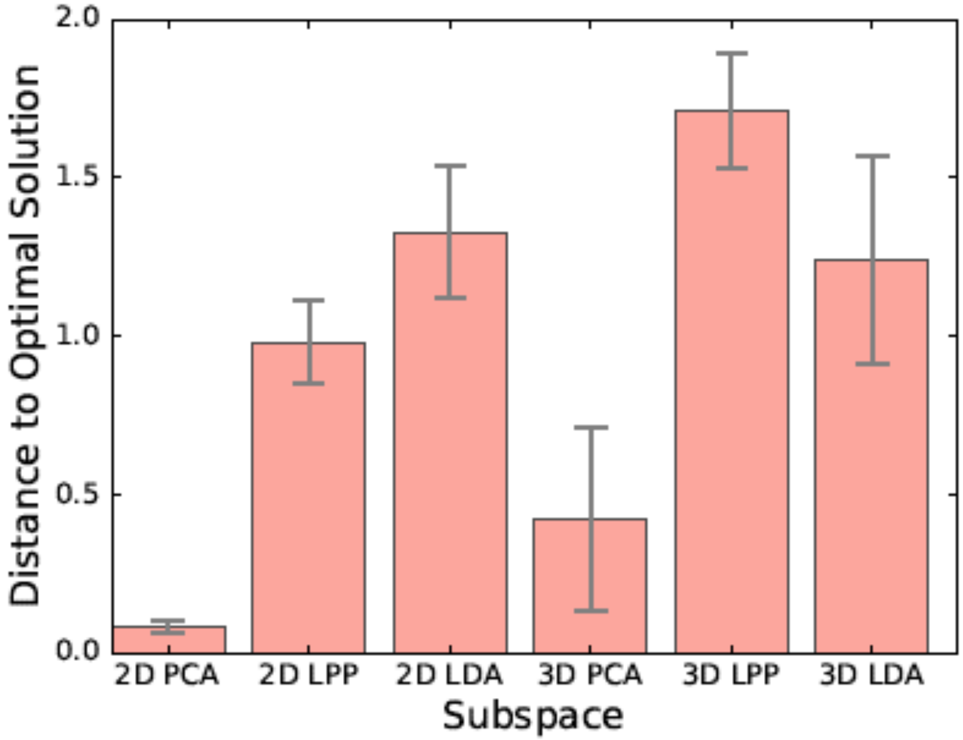
Ecoli Dataset (7 dimensional)





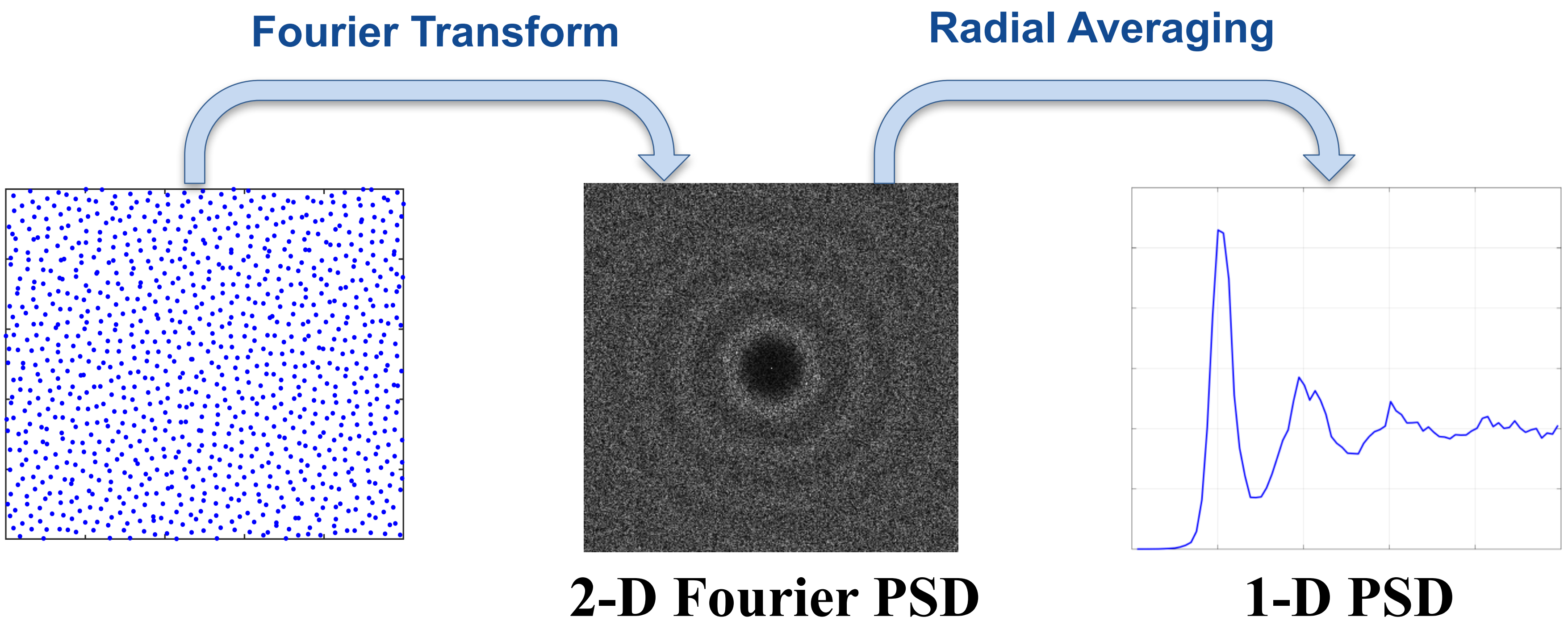
Diverse Solutions

- Grassmann distance between the optimal solution and the linear subspace inferred using the consensus from PDS samples for the Ecoli dataset.



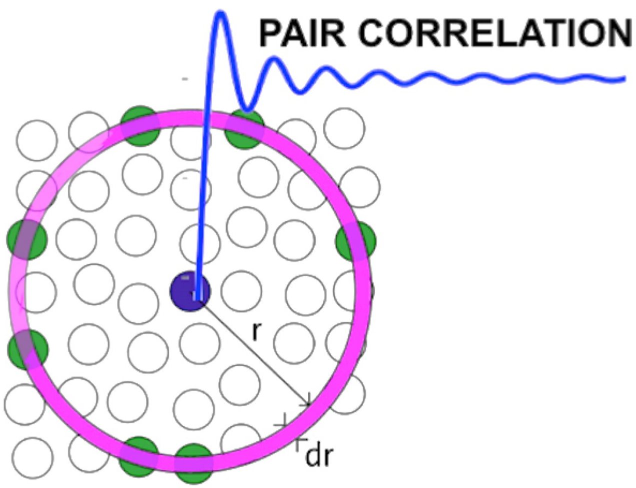
Why Does PDS Work?

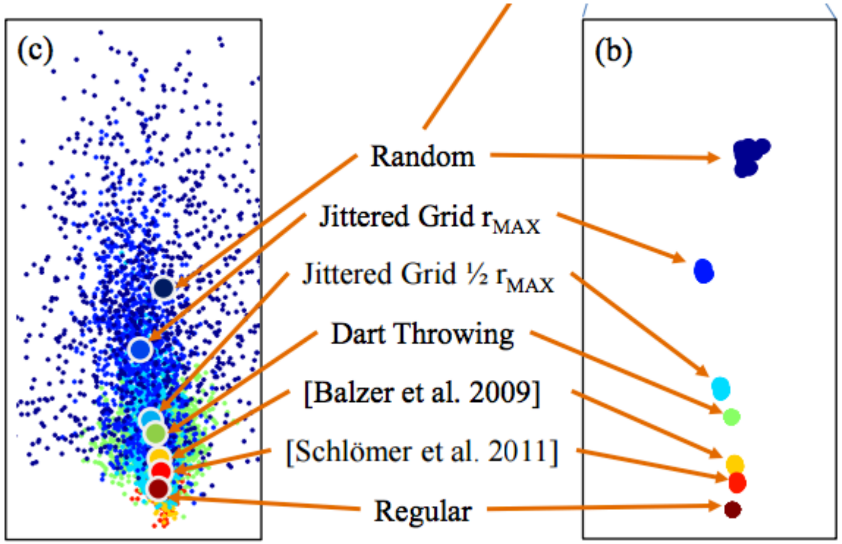
- Sample design can be directly viewed from a ***function reconstruction*** perspective – Fourier Domain Analysis



Characterizing Spatial Statistics

- In statistical mechanics, Pair Correlation Function describes how density varies as a function of distance from a reference

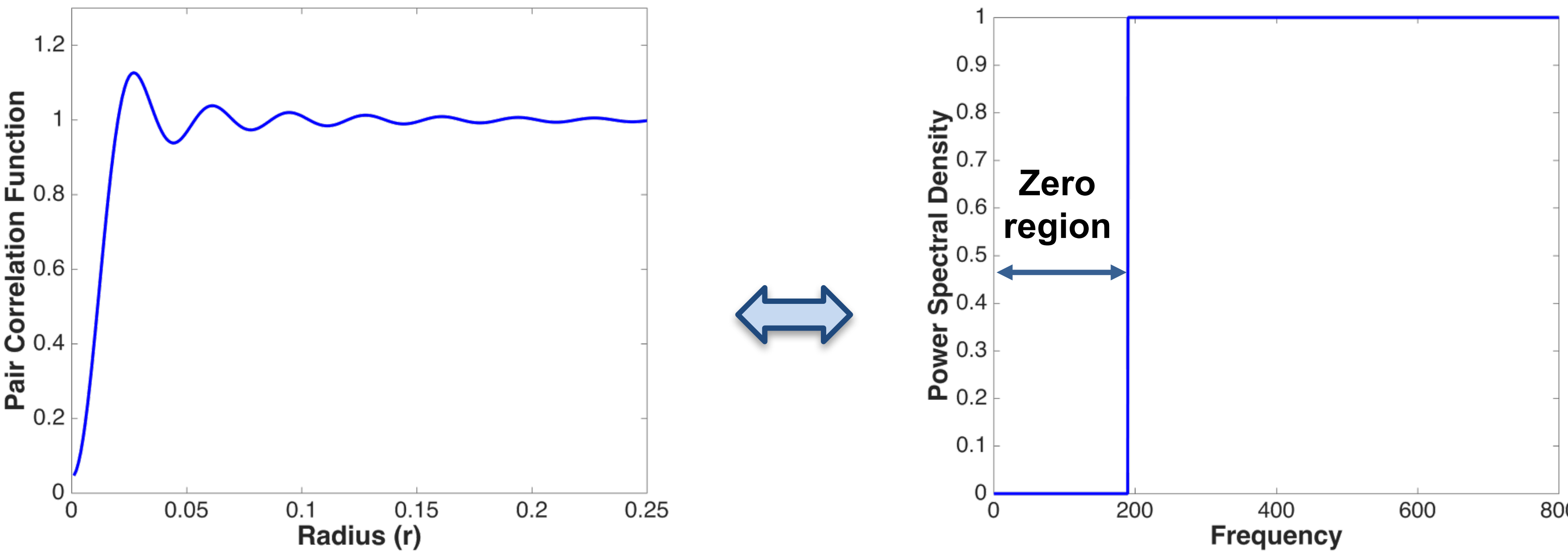




Key Idea for Spectral Sampling

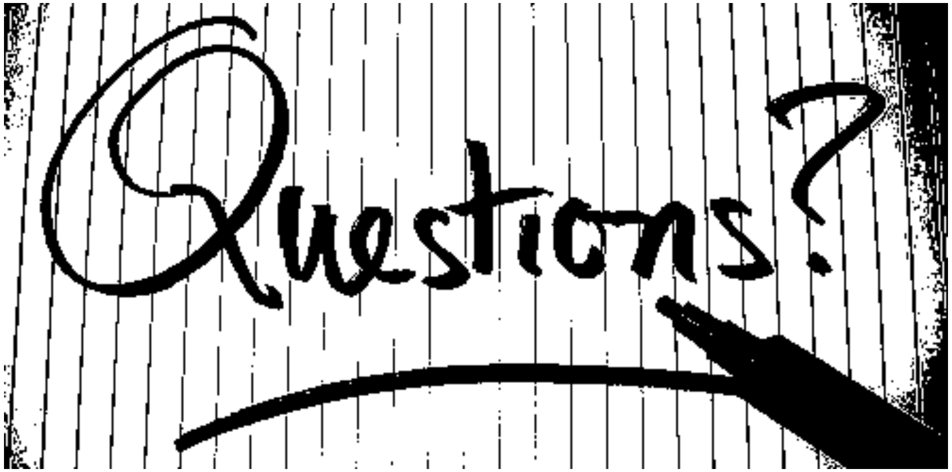
- 1-D PSD and PCF are connected via the *1-D* Hankel transform

$$P(k) = 1 + \rho(2\pi)^{\frac{d}{2}} k^{1-\frac{d}{2}} H_{\frac{d}{2}-1}\left(r^{\frac{d}{2}-1}(G(r)-1)\right)$$



Summary

- Poisson disk sampling for better coverage of Grassmann
- Effective approximation of the optimal solution for subspace learning problems
- High-quality samples can serve as anchor points to search through the Grassmannian using conventional optimization strategies



Contact

Bhavya Kailkhura
Center for Applied Scientific Computing
Lawrence Livermore National Laboratory
Email: kailkhura1@llnl.gov